VU Research Portal

Reproducibility of isometric trunk extension torque, trunk extensor endurance, and related electromyographic parameters in the context of their clinical applicability

Van Dieen, J. H.; Heijblom, P.

published in
Journal of Orthopaedic Research
1996

DOI (link to publisher)
10.1002/jor.1100140122

Link to publication in VU Research Portal

citation for published version (APA)
Van Dieen, J. H., & Heijblom, P. (1996). Reproducibility of isometric trunk extension torque, trunk extensor endurance, and related electromyographic parameters in the context of their clinical applicability. *Journal of Orthopaedic Research*, *14*(1), 139-143. https://doi.org/10.1002/jor.1100140122

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:
vuresearchportal.ub@vu.nl

Download date: 04. May. 2023

Journal of Orthopaedic Research
14:139-143 The Journal of Bone and Joint Surgery, Inc.
© 1996 Orthopaedic Research Society

Reproducibility of Isometric Trunk Extension Torque, Trunk Extensor Endurance, and Related Electromyographic Parameters in the Context of Their Clinical Applicability

*†Jaap H. van Dieën and *P. Heijblom

**Faculty of Human Movement Sciences, Vrije Universiteit, Amsterdam, and †Work Study and Ergonomics Department, IMAG-DLO, Wageningen, The Netherlands*

Summary: Testing the capacity of the trunk extensor muscles may be useful in the diagnosis of low back pain. In the present study, the reproducibility of measurements of maximum trunk extension force, trunk extension endurance, and related electromyographic parameters was investigated. Intraclass correlations indicated that the reproducibility of maximum force and endurance time was satisfactory. Nevertheless, the smallest difference in these parameters that could be attributed, with 95% confidence, to a change in the condition of a patient was, in general, more than 20%. On the electromyograms, the slopes of amplitude and frequency content appeared to be related to endurance time. The reproducibility of these parameters in terms of the intraclass correlation was again satisfactory; however, the smallest detectable difference generally exceeded 50%. The clinical applicability of the parameters studied is severely limited by a lack of reproducibility.

Reduced capacity of force or endurance in the back muscles has been reported to be related to low back pain, either as a cause or as a consequence of the disorder (2,8,9,14). Therefore, tests of maximum force or endurance may be useful in the diagnosis of low back pain. The clinical applicability of such tests is limited, however, because pain and motivation felt by the subject may have a strong influence on the results. Therefore, the development of more objective and preferably less stressful tests, possibly based on the use of electromyography, would seem desirable. The rates of change for electromyographic parameters, as described by estimates of the slope of their time-series, are related to endurance time (7,16). These electromyographic changes can be used to discriminate between patients with low back pain and healthy individuals (8,11,12,15). Klein et al. (6) showed that such electromyographic parameters discriminate between patients and healthy subjects even better than more traditional measurements, such as maximum force and range of motion.

If maximum force, endurance, or electromyographic parameters are to be used as tools in diagnostics or in monitoring therapeutic intervention, a high degree of reproducibility is demanded. The few studies that have dealt with this topic usually have expressed reproducibility by the coefficient of intraclass correlation. However, the interpretation of the intraclass correlation, with respect to the clinical application of a parameter, is not straightforward. When a parameter is used to monitor a subject's status in time, more information can be obtained from the smallest detectable difference (10,13). The smallest detectable difference reflects the percentage change in a parameter that can be attributed with 95% certainty to a real change in the condition of the subject instead of being caused by test-retest errors. The present study involved two experiments in which the intraclass correlation coefficient and the smallest detectable difference were determined for the maximum force, the endurance time, and the rates of change in the electromyographic parameters of the trunk extensors.

METHODS

Subjects

To obtain sufficient variability between subjects, the first experiment included groups of trained and untrained subjects. The trained group consisted of five varsity rowers. They each had 2-7 years of rowing experience and trained about 10 hours per week. Their mean (±SD) age, height, and mass were 22 ± 3 years, 1.92 ± 0.05 m, and 85 ± 9 kg. The untrained group consisted of five men who did not participate in any regular sports activity or in any other activity involving forceful contractions of the low back muscles. Their mean age, height, and mass were 27 ± 8 years, 1.85 ± 0.07 m, and 77 ± 3 kg.

The second experiment involved a group of nine men who were not selected specifically with respect to the training status of the

Received February 3, 1995; accepted May 30, 1995.
Address correspondence and reprint requests to J. H. van Dieën at IMAG-DLO, P.O. Box 43, NL-6700 AA Wageningen, The Netherlands.

back muscles. Their mean age, height, and mass were 24 ± 3 years, 1.83 ± 0.05 m, and 74 ± 4 kg. None of the subjects reported any low back pain at the time of the experiment.

Instrumentation

Both experiments were performed with the subjects seated on a dynamometer (Biodex Medical Systems, Shirley, NY, U.S.A.). The torque exerted on the axis of the dynamometer was digitized at 50 Hz by a DAS8 AD-converter (Keithly Metrabyte, Taunton, MA, U.S.A.), stored on the hard disk of a personal computer, and fed back to the subject on a video display unit.

In the second experiment only, the electromyographic signals from six electrode locations were recorded. Electrodes were placed longitudinally over the left and right multifidus, iliocostalis lumborum, and longissimus thoracis muscles, as described previously (16). The signals were preamplified and transmitted to a Biomes 8 receiver (Glonner Electronic, Munich, Germany) and recorded on tape (SR-70; TEAC, Tokyo, Japan).

Protocol

Experiment 1 consisted of two sessions, each involving two trials. At the start of each session, the subject's maximum voluntary contraction torque was determined according to the procedure proposed by Caldwell et al. (3). After 5 minutes of rest, the subject performed a sustained contraction at 50% of the maximum. When, after repeated encouragement, the torque no longer could be maintained above 90% of the target level, the test was stopped and endurance time (T_{end}) was recorded. After 30 minutes of rest, the procedure was repeated with the endurance test at the same absolute target level. To exclude variation due to repositioning, subjects were not allowed to step from the dynamometer during the rest period. The second session was performed several days later.

In the second experiment, the subject performed a warm-up consisting of 30 2-second isokinetic contractions at 100 Nm. The subject then rested for 5 minutes and subsequently performed seven forceful (near maximum) isometric contractions to practice producing a maximum contraction. After 5 minutes of rest, the maximum voluntary contraction torque was determined. Subsequently, the subject again rested for 5 minutes and then sustained a contraction at 80% of the maximum until the limit of endurance was reached. The entire session was repeated at least 2 days later.

Data Analysis

The electromyographic signals obtained in the second experiment were replayed from tape and digitized at 1,024 Hz. All signals were inspected visually to control signal quality. The rectified and averaged amplitude of the electromyogram was determined over each 3 seconds of T_{end}. The mean power frequency in each 3-second sample was determined by calculating and averaging the spectra of 80% overlapping windows of 1,024 samples with use of a fast Fourier transformation. Linear regression analysis was used to determine the slopes of the electromyographic parameters versus time. The slopes also were determined for the two parameters after they were normalized to their initial values.

Only slopes deviating significantly from zero were used in the subsequent analysis. The relationship between the slope of the electromyographic parameters and $\log(T_{end})$ was calculated by linear regression analysis. The reproducibility of the maximum voluntary contraction torque, T_{end}, and the electromyographic parameters was expressed by means of the intraclass correlation coefficient, i.e., the ratio of the intersubject variation and the total variation in the data set. In addition, the smallest detectable difference was calculated. The smallest detectable difference is equal to a constant ($1.96\sqrt{2}$), multiplied by the square root of the summed components of the variance due to test-retest errors and their interaction with the subject, expressed as a percentage of the grand mean of the parameter.

TABLE 1. *Reproducibility of maximum extension torque and endurance time for experiment 1*

	ICC	SDD (%)
MVC		
Within session 1	0.71	26.6
Within session 2	0.97	9.8
Trial 1, between sessions	0.84	20.5
Trial 2, between sessions	0.84	21.7
T_{end}		
Within session 1	0.94	29.3
Within session 2	0.95	23.1
Trial 1, between sessions	0.54	73.5
Trial 2, between sessions	0.37	92.4
$\log(T_{end})$/MVC		
Within session 1	0.82	30.1
Within session 2	0.99	9.1
Trial 1, between sessions	0.77	37.9
Trial 2, between sessions	0.94	20.7

The reproducibility of the values for maximum extension torque (MVC) and endurance time at 50% of the MVC (T_{end}), as expressed by the intraclass correlation coefficient (ICC) and the smallest detectable difference (SDD).

RESULTS

Experiment 1

The first experiment was aimed at determining the reproducibility of maximum voluntary contraction torque and T_{end}. Analysis of variance revealed a significantly higher mean (±SD) maximum contraction torque for the trained group (359 ± 23 compared with 269 ± 35 Nm), although T_{end} was significantly lower in this group (102 ± 29 compared with 173 ± 57 seconds). Since maximum contraction torque has been shown to have an effect on T_{end}, a multiple regression analysis of maximum voluntary contraction and training on $\log(T_{end})$ was performed. The maximum voluntary contraction appeared to account for 33-62% of the variance in T_{end}, and no independent effect of training was found. A paired *t* test revealed no significant differences between trials and sessions.

The reproducibility of the maximum voluntary contraction torque and T_{end} was described by the intraclass correlation and the smallest detectable difference (Table 1). For this analysis, the data from both groups were pooled. The reproducibility of the maximum voluntary contraction torque was quite satisfactory, although the magnitude of the smallest detectable difference was still considerable. Between sessions, the reproducibility of T_{end} was clearly not as good. Of course, test-retest variation in the maximum voluntary contraction torque led to variation in the absolute force level and as such it contributed to variation in T_{end}. This is illustrated by the higher reproducibility achieved within sessions, when one absolute force level per subject was used. In view of the relationship between T_{end} and

TABLE 2. *Coefficient of correlation for the logarithm of the endurance time with the slopes of the electromyographic parameters for experiment 2*

	RAEMG		RAEMGn		MPF		MPFn	
	Session 1	Session 2	Session 1	Session 2	Session 1	Session 2	Session 1	Session 2
Left longissimus	−0.75	−0.73	NS	NS	0.90	0.68	0.87	0.71
Right longissimus	−0.77	−0.77	−0.64	−0.65	0.80	NS	0.80	NS
Left multifidus	−0.84	−0.80	NS	−0.81	0.78	0.76	0.84	0.70
Right multifidus	−0.94	NS	−0.74	NS	0.86	0.63	0.90	0.88
Left iliocostalis	NS	−0.85	−0.90	NS	NS	NS	NS	NS
Right iliocostalis	−0.77	−0.81	−0.74	−0.71	NS	0.80	NS	0.77

The logarithm of the endurance time ($\log[T_{end}]$) was correlated with the slopes of the rectified and averaged electromyogram (RAEMG) and the mean power frequency (MPF), as well as with these parameters normalized to their initial values (RAEMGn and MPFn). NS = nonsignificant.

maximum voluntary contraction, it could be argued that in order to discriminate between patients with low back pain and healthy subjects, T_{end} should be corrected for the maximum voluntary contraction torque. Since this relationship is curvilinear, the ratio between $\log(T_{end})$ and maximum voluntary contraction torque was calculated. This procedure also corrected for errors in the determination of the maximum voluntary contraction torque. The reproducibility of this parameter was superior to the reproducibility of T_{end} alone (Table 1).

Experiment 2

The mean (±SD) maximum voluntary contraction torque in experiment 2 was 390 ± 97 Nm in the first session and 393 ± 90 Nm in the second session. The mean (±SD) T_{end} at 80% of the maximum voluntary contraction torque in the first session was 44 ± 18 seconds and in the second session it was 46 ± 19 seconds. No significant differences for T_{end} existed between sessions. The intraclass correlation coefficients for maximum voluntary contraction and T_{end} were, respectively, 0.93 and 0.68, with values for smallest detectable difference of 18 and 67%. The ratio of $\log(T_{end})$ and the maximum voluntary contraction had an intraclass correlation coefficient of 0.94 with a smallest detectable difference of 24%.

After the quality of the electromyographic signals was controlled, the data from two electrode locations in each of two subjects for one session were discarded. For another subject, the data from one location for one session were discarded. For the remaining 103 time-series, a linear fit was used to estimate the slope of the electromyographic parameters. The slope of the rectified and averaged amplitude did not appear to be significantly different from zero in 12 of these samples, of which nine were obtained from one subject. Of the slopes for the mean power frequency, 15 were not significantly different from zero, and seven of the 15 were obtained from one subject with a very low T_{end}. The nonsignificant slope values were not analyzed further, leaving 91 values for the rectified and averaged amplitude and 88 for the mean power frequency.

Scatterplots of electromyographic parameters versus T_{end} indicated curvilinear relationships with nonhomogeneous distributions of the residuals. Linear regression of $\log(T_{end})$ on each electromyographic parameter was used to describe these relationships. Table 2 gives an overview of the results. Normalization enhances the relationship for the mean power frequency, whereas it lowers the results for the rectified and averaged amplitude. Hence, the slopes of the rectified and averaged amplitude and the normalized mean power frequency appear to be the most useful of the parameters studied. The reproducibility of these two parameters was evaluated by the intraclass correlation and the smallest detectable difference. The results are presented in Table 3. On average, the reproducibility of the slope of the normalized mean power frequency was superior to the slope of the rectified and averaged amplitude. Although the intraclass correlation coefficients tended to be high in general, the values for the smallest detectable difference were

TABLE 3. *Reproducibility of the slopes of the electromyographic parameters for experiment 2*

	RAEMG slope		MPFn slope	
	ICC	SDD (%)	ICC	SDD (%)
Left longissimus	0.91	61	0.90	65
Right longissimus	0.81	112	0.84	71
Left multifidus	0.68	104	0.87	66
Right multifidus	0.31	188	0.88	75
Left iliocostalis	0.58	142	0.91	25
Right iliocostalis	0.84	105	0.86	51

The reproducibility of the values for the rectified and averaged electromyogram (RAEMG) and the normalized mean power frequency (MPFn), as expressed by the intraclass correlation coefficient (ICC) and the smallest detectable difference (SDD).

over 50%. In contrast to the results for T_{end}, the reproducibility of the electromyographic parameters did not change substantially when corrected for intersession variance in the maximum voluntary contraction.

DISCUSSION

Experiment 1 shows, in accordance with previous work (14), that the maximum contraction torque for isometric trunk extension in the upright position can be determined quite reproducibly. However, as shown by the values for the smallest detectable difference, test-retest errors between sessions may be as large as 20%. The higher intraclass correlation coefficient in session 2 was probably caused by the subjects being more accustomed to the test. T_{end} was not very reproducible across sessions. Correcting T_{end} for the maximum voluntary contraction torque clearly improved this. Within sessions, reproducibility was much better when endurance was tested at the same absolute force level. In line with this, a high intraclass correlation coefficient (0.82-0.99) of T_{end} for trunk extension at a constant absolute force level has been reported (5,7). Thus, the low reliability of T_{end} is caused in part by test-retest variation in the results for maximum voluntary contraction. This led to the decision to use the longer warm-up period and the training for forceful contractions in the second protocol. Obviously, the use of this option is limited in clinical applications. Fuglevand et al. (4) have shown the reliability of T_{end} to increase with relative force level. Therefore, a higher relative force (80 compared with 50% maximum voluntary contraction) was used also. The changes resulted in somewhat better reproducibility of the values for maximum voluntary contraction. The reproducibility of T_{end} was also better, although still not satisfactory. Correction for the maximum voluntary contraction again increased reliability considerably.

Experiment 1 demonstrates the limitations of intraclass correlation. Whereas intraclass correlation coefficient values greater than 0.80 generally are considered satisfactory, the smallest detectable difference still may be as large as 30%. This shows that the reliability of a parameter to be used in follow-up measurements on a subject cannot be evaluated adequately from the intraclass correlation. In addition, when the degree of reproducibility is determined for a study on a homogeneous group, low values for the intraclass correlation coefficient may correspond to a low smallest detectable difference. For instance, the intraclass correlation coefficient of the maximum voluntary contraction of the trained subjects in experiment 1 was as low as 0.03, whereas the smallest detectable difference was only 25%. Therefore, intersubject variation in a group participating in an experiment to determine reproducibility should be representative of the population to which the method will be applied. It should be noted in this respect that the intersubject coefficient of variation of T_{end} was similar in experiments 1 and 2, although subjects in experiment 1 were selected specifically to obtain a high variance.

As in previous studies (7,16), curvilinear relationships were found between the rate of change of the electromyographic parameters and the endurance time of the trunk extensor muscles. The intraclass correlation of the electromyographic parameters showing the strongest relationships overall was quite satisfactory, although, especially for the slope of the rectified and averaged amplitude, the smallest detectable difference was high. In a similar experiment, Roy et al. (11) found a somewhat higher reliability (intraclass correlation coefficient = 0.94). However, the tests were performed only 15 minutes apart and the electrodes were not replaced. In addition, data from six different electrode locations on the same subject were pooled and treated as statistically independent. This would tend to inflate the intraclass correlation. Mannion and Dolan (7) found an intraclass correlation coefficient of 0.98 for the slope of the median frequency of the lumbar trunk extensors. These tests were performed 1-2 weeks apart and the electrodes were replaced. However, the number of subjects (n = 5) was fairly small, which hampers interpretation since the intraclass correlation coefficient depends on the intersubject variation of the parameter studied.

Biedermann (1) has argued that, if fatigue effects are relatively small in a group of subjects and if the error of the estimates is constant, then the intraclass correlation coefficient for the electromyographic parameters would be rather low. This would be expected to lead to a lower intraclass correlation in healthy subjects as compared with patients with low back pain, as was confirmed by experimental data (1). When the confidence intervals of the retest values were expressed in percentages of the test value, a procedure comparable with the calculation of the smallest detectable difference, the difference between the groups disappeared. Since it is questionable whether errors are constant rather than proportional to the mean of the estimates, this might also be explained by lower intersubject variation in the group of healthy controls, as was demonstrated above. In addition, at 80% of the maximum voluntary contraction torque, the force level used in the present study, fatigue effects are pronounced in healthy subjects as well.

More importantly, Biedermann (1) has argued that considerable test-retest variation in the electromyographic parameters is acceptable, since the metabolic status of the muscles, which they are thought to reflect, can be expected to be equally variable. Although this would indicate that the metabolic status of the muscle can be screened with validity by means of electromyographic parameters, such high variability would limit the clinical value of this procedure severely.

The reliability of T_{end} greatly improved when corrected for the differences in maximum voluntary contraction torque. This did not hold true for the electromyographic parameters; their reliability remained more or less constant. However, a similar correction might be useful in diagnostic applications; for instance, in avoiding misclassification of trained subjects (11).

The results of the present study, especially regarding the smallest detectable difference, show that the utility of these parameters in monitoring patients with low back pain is hampered severely by the limited reliability of the parameters. The estimation of the smallest detectable difference is based on partitioning of the variation in the data set into components related to inter and intrasubject variability. The confidence intervals of these variation components are rather wide, causing unstable estimates of the smallest detectable difference (13). However, if one ignores the precise numerical outcomes, the present results imply that only large changes in T_{end} and related parameters can be detected reliably. Therefore, it appears necessary to use multiple criteria to monitor a patient's condition, as was also implied for diagnosis (8,11). It is possible that more reliable parameters can be designed by scaling data from various muscles.

Acknowledgment: We would like to express our gratitude to our students Victor Plomp, Coen van den Bemt, Bent Bervoets, and Brian Wongsopawiro for their assistance in collecting and analyzing the data, and to Dr. Michiel de Looze for his useful comments on the manuscript. Furthermore, the authors wish to acknowledge Hoogovens bv. for permitting us to use their Biodex dynamometer.

REFERENCES

1. Biedermann HJ: A method for assessing the equivalence of repeated measures of muscle fatigue rates estimated from electromyography power spectrum analysis. *J Electromyogr Kinesiol* 1:288-292, 1991
2. Biering-Sørensen F: Physical measurements as risk indicators for low-back trouble over a one-year period. *Spine* 9:106-119, 1984
3. Caldwell LS, Chaffin DB, Dukes-Dobos FN, Kroemer KHE, Laubach LL, Snook SH, Wasserman DE: A proposed standard procedure for static muscle strength testing. *Am Ind Hyg Assoc J* 35:201-206, 1974
4. Fuglevand AJ, Zackowski KM, Huey KA, Enoka RM: Impairment of neuromuscular propagation during human fatiguing contractions at submaximal forces. *J Physiol* 460:549-572, 1993
5. Jørgensen K, Nicolaisen T: Two methods for determining trunk extensor endurance: a comparative study. *Eur J Appl Physiol* 55:639-644, 1986
6. Klein AB, Snyder-Mackler L, Roy SH, De Luca CJ: Comparison of spinal mobility and isometric trunk extensor forces with electromyographic spectral analysis in identifying low back pain. *Phys Ther* 71:445-454, 1991
7. Mannion AF, Dolan P: Electromyographic median frequency changes during isometric contraction of the back extensors to fatigue. *Spine* 19:1223-1229, 1994
8. Mayer TG, Kondraske G, Mooney V, Carmichael TW, Butsch R: Lumbar myoelectric spectral analysis for endurance assessment: a comparison of normals with deconditioned patients. *Spine* 14:986-991, 1989
9. Nicolaisen T, Jørgensen K: Trunk strength, back muscle endurance and low-back trouble. *Scand J Rehabil Med* 17:121-127, 1985
10. Roebroeck ME, Harlaar J, Lankhorst GJ: The application of generalizability theory to reliability assessment: an illustration using isometric force measurements. *Phys Ther* 73:386-401, 1993
11. Roy SH, De Luca CJ, Casavant DA: Lumbar muscle fatigue and chronic lower back pain. *Spine* 14:992-1001, 1989
12. Roy SH, De Luca CJ, Snyder-Mackler L, Emley MS, Crenshaw RL, Lyons JP: Fatigue, recovery, and low back pain in varsity rowers. *Med Sci Sports Exerc* 22:463-469, 1990
13. Shavelson RJ, Webb NM: *Generalizability Theory: A Primer.* Newbury Park, California, Sage Publications, 1991
14. Smidt G, Herring T, Amundsen L, Rogers M, Russell A, Lehmann T: Assessment of abdominal and back extensor function: a quantitative approach and results for chronic low-back patients. *Spine* 8:211-219, 1983
15. Tsuboi T, Satou T, Egawa K, Izumi Y, Miyazaki M: Spectral analysis of electromyogram in lumbar muscles: fatigue induced endurance contraction. *Eur J Appl Physiol* 69:361-366, 1994
16. van Dieën JH, Oude Vrielink HHE, Housheer AF, Lötters FBJ, Toussaint HM: Trunk extensor endurance and its relationship to electromyogram parameters. *Eur J Appl Physiol* 66:388-396, 1993